



Chinese Society of Aeronautics and Astronautics
& Beihang University

Chinese Journal of Aeronautics

cja@buaa.edu.cn
www.sciencedirect.com



Impulse feature extraction method for machinery fault detection using fusion sparse coding and online dictionary learning



Deng Sen ^{a,b,*}, Jing Bo ^a, Sheng Sheng ^a, Huang Yifeng ^a, Zhou Hongliang ^a

^a Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China

^b Unit 94371 of People's Liberation Army, Zhengzhou 450045, China

Received 19 May 2014; revised 7 August 2014; accepted 31 October 2014

Available online 21 February 2015

KEYWORDS

Dictionary learning;
Fault detection;
Impulse feature extraction;
Information fusion;
Sparse coding

Abstract Impulse components in vibration signals are important fault features of complex machines. Sparse coding (SC) algorithm has been introduced as an impulse feature extraction method, but it could not guarantee a satisfactory performance in processing vibration signals with heavy background noises. In this paper, a method based on fusion sparse coding (FSC) and online dictionary learning is proposed to extract impulses efficiently. Firstly, fusion scheme of different sparse coding algorithms is presented to ensure higher reconstruction accuracy. Then, an improved online dictionary learning method using FSC scheme is established to obtain redundant dictionary and it can capture specific features of training samples and reconstruct the sparse approximation of vibration signals. Simulation shows that this method has a good performance in solving sparse coefficients and training redundant dictionary compared with other methods. Lastly, the proposed method is further applied to processing aircraft engine rotor vibration signals. Compared with other feature extraction approaches, our method can extract impulse features accurately and efficiently from heavy noisy vibration signal, which has significant supports for machinery fault detection and diagnosis.

© 2015 The Authors. Production and hosting by Elsevier Ltd. on behalf of CSAA & BUAA. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Fault feature extraction plays a key role in detecting failures of machinery engineering systems such as aircraft engine, space shuttle and rotating machinery. In recent decades, signal processing theory has been widely used in feature extraction for fault diagnosis. Periodical impulse components in vibration signals are important indicators of system health status. Therefore, many advanced signal processing methods have been studied to extract the impulse components from vibration

* Corresponding author at: Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China. Tel.: +86 29 84787628.

E-mail address: 425931056@qq.com (S. Deng).

Peer review under responsibility of Editorial Committee of CJA.



Production and hosting by Elsevier

signals.¹ However, due to the factors of sensor inaccuracy, operator error and electromagnetic interference, the measured vibration signals may contain different kinds of noises. How to extract impulse components effectively from noisy vibration signals can be more attractive.^{2,3}

Wavelet analysis is considered to be effective for fault features extraction from vibration signals. The wavelet shrinkage method proposed by Donoho achieved huge success in removing noises and extracting features.⁴ Qiu et al. presented wavelet filtering to detect periodical impulse components from vibration signals.⁵ He et al. proposed a hybrid method comprised of wavelet filter and morphological processing to get weak mechanical impulse.⁶ However, when the noise strength is too high, wavelet analysis methods may reduce signal energy and discard important impulse features during the process of removing noises. Moreover, without the prior information of noises in vibration signal, it is difficult to select wavelet filters' parameters which have great influence on the performance of feature extraction.

Sparse coding (SC) is a new signal processing method which has found many applications in the solution of image processing,⁷ signal de-noising,⁸ compressed sensing,⁹ etc. The noisy vibration signal can be most sparsely represented in sparse domain by redundant dictionary, which provides an efficient way to reconstruct input signal and extract impulse components. The sparse coefficients of noises are zero or nearly zero, but impulse components have large sparse coefficients. Thus, the noises in vibration signal are not reconstructed in the process of sparse reconstruction and the sparse signal obtained by using SC method is the estimation of noiseless impulse components. The key issues of extracting features based on SC method include sparse reconstruction and redundant dictionary selection. Many researchers have focused on this field, and recent contributions for extracting impulse features using sparse model are proposed.^{10–12} Liu et al. proposed a shift-invariant sparse coding (SISC) algorithm to get basis functions of redundant dictionary separately and extract sparse features for fault diagnosis.¹⁰ Tang et al. also used SISC algorithm to learn redundant dictionary and an optimal latent components' filtering algorithm was designed to extract features based on this redundant dictionary.¹¹ Chen et al. presented a new scheme called sparse extraction of impulse by adaptive dictionary (SpaEIAD) to represent vibration signal and extract impulse components.¹² The problem of dictionary selection is to choose a fixed dictionary such as the wavelet basis and Fourier basis, or learn a redundant dictionary adapted to the input signal. Dictionary learning method achieves sparse signal representation by learning redundant dictionary from training samples, which can match the structure of input signal. Typical learning algorithms include the K-SVD,¹³ the sparse K-SVD,¹⁴ the method of direction (MOD),¹⁵ etc. In practice, vibration signal often contains mass data samples. The above dictionary learning methods cannot handle large training sets effectively because the whole training sets are used to solve a constrained optimization problem at each iteration. Moreover, classical sparse reconstruction algorithms include basis pursuit (BP), orthogonal matching pursuit (OMP), subspace pursuit (SP), etc, and it is difficult to select an appropriate sparse reconstruction algorithm which has great effect on the performance of extracting features and recovering signals.

Generally, it is essential to design appropriate sparse coding and dictionary learning methods for extracting impulse

features. In data fusion theory, information collected from different sensors can be fused to get higher reliable information. Fusion scheme of several estimators has been studied to form a better estimator.^{16,17} Ambat et al. presented a fusion compressed sensing algorithm to improve the signal reconstruction performance.¹⁸ If the estimated results of various SC algorithms are collected and fused effectively, the accuracy of recovering signals will be improved remarkably. So a fusion scheme of SC algorithms is designed in this paper. To deal with large training samples rapidly, Mairal et al. proposed a new online dictionary learning method based on stochastic approximation algorithm.¹⁹ The online approach processed one element of the training set at a time, which is suitable to learning large vibration data samples with low computational time. But this online method in Ref.¹⁹ only used singular SC algorithm (Lasso) and it would affect the accuracy of solving sparse coefficients and learning redundant dictionary. If fusion scheme of SC algorithm is used to solve the sparse coefficients in online dictionary learning method, the redundant dictionary will be learned accurately to adapt to the input signals. The advantages of fusion sparse coding (FSC) algorithm and improved online dictionary learning method are beneficial to extracting impulse features from vibration signals accurately and effectively, especially for the vibration signal with heavy background noises. Commonly, the noise power of heavy noisy vibration signal is large and the sparse coefficients of noises solved by some singular SC algorithms may be far greater than zero. The inaccuracy of solving sparse coefficients causes the learned redundant dictionary not adapt to the structure of input signal. FSC and improved online learning method can solve the sparse coefficients and update the redundant dictionary more accurately and effectively, which can significantly improve the performance of sparse reconstruction and dictionary learning for extracting impulse features from heavy noisy vibration signal.

This paper is organized as follows: In Section 2, the fundamentals about sparse coding and dictionary learning are introduced. Section 3 presents FSC algorithm and improved online dictionary learning method in detail. In Section 4, the scheme of extracting impulse features using the proposed method is presented, and the performance of this method is validated through simulations and comparisons. In Section 5, the proposed method is applied to processing aircraft engine rotor vibration signals with heavy background noises, which can illustrate the capability of this method to identify the impulse features for machinery fault diagnosis. Finally, conclusions and summaries are presented in Section 6.

2. Sparse coding and redundant dictionary design

Sparse coding is described as a generative model that an input signal can be represented as a linear combination of basis functions with additive noise. Denote a measured noisy signal $\mathbf{x} \in \mathbf{R}^p$ and noiseless signal $\mathbf{z} \in \mathbf{R}^p$, the model can be represented as

$$\mathbf{x} = \mathbf{z} + \mathbf{w} = \mathbf{D}\mathbf{s} + \mathbf{w} \quad (1)$$

where $\mathbf{D} \in \mathbf{R}^{p \times n}$ is called redundant dictionary, and the dictionary consists of n basis functions $\mathbf{d}^j \in \mathbf{R}^p (j = 1, 2, \dots, n)$. $\mathbf{s} \in \mathbf{R}^n$ called sparse coefficients of the input signal \mathbf{x} , and $\mathbf{w} \in \mathbf{R}^p$ an additive zeros-mean Gaussian noise with variance

σ^2 . In this model, the number of basis functions should be greater than the dimension of input signal with $n > p$.

Sparse coding is the process of determining sparse coefficients s using the redundant dictionary D which is specified manually or learned from input signal x . Eq. (1) has infinite solutions because the redundant dictionary D has greater columns than its rows. However, most of the elements in s are zero or nearly zero, which mean input signal x can be recovered using small number of basis functions in dictionary D . The sparsest solution in Eq. (1) can be solved by the following optimization problem

$$\min_s \|s\|_0 \quad \text{s.t.} \quad \|x - Ds\|_2 \leq \gamma \quad (2)$$

where $\|\cdot\|_0$ and $\|\cdot\|_2$ denote l_0 - and l_2 - norm, respectively. Parameter γ is the approximation error tolerance. The redundant dictionary D can be assigned as specified transform matrix or learned from training data adapted to the structure of signals. The sparse coefficients s can be estimated using Eq. (2) with D and x fixed:

$$\hat{s} = \underset{s}{\operatorname{argmin}} \|x - Ds\|_2^2 + \mu \|s\|_1 \quad (3)$$

Solving the sparse coefficients s in Eq.(2) proves to be a NP-hard problem because the l_0 - norm sparsest solution is a non-convex optimization problem. Some algorithms based on greedy strategy have been used to solve the problem such as matching pursuit (MP), OMP, etc, which can compute coefficients s sequentially. Other algorithms solve sparse coefficients based on maximum a posterior (MAP) estimation theory and more details can be found in Refs.^{20–22}. The MAP method estimates coefficients s as random variables by maximizing the posterior likelihood function, which means the problem can be solved through convex optimization because the l_0 - norm in Eq. (2) is replaced of the l_1 - norm in Eq. (3). These methods include BP,²³ Lasso,²⁴ coordinate descent (CD),²⁵ etc. Similar algorithm is focal underdetermined system solver (FOCUSS) that uses l_p - norm as a replacement of l_0 - norm.²⁶ Some advanced sparse reconstruction algorithms have also been proposed in order to seek the sparse coefficients s .^{27,28}

Designing redundant dictionary D is another important issue for sparse signal reconstruction. For certain special signal, it is difficult to select an appropriate redundant dictionary to match the structure of input signal. Dictionary learning method can obtain a proper dictionary D adapted to the input signal, and the method learns redundant dictionary based on maximum likelihood (ML) estimation theory.^{13,14,20} Given the training set $X = \{x_i\}_{i=1}^N$, the dictionary D can be learned by the following joint optimization problem:

$$\hat{D} = \underset{D}{\operatorname{argmin}} \sum_{i=1}^N \min_{s_i} \{\|Ds_i - x_i\|_2^2 + \mu \|s_i\|_1\} \quad (4)$$

where s_i denotes sparse coefficients corresponding to the training samples x_i . μ a regularization parameter. Approach to solve the joint optimization problem in Eq. (4) involves two sub-problems: solving sparse coefficients s and computing redundant dictionary D , whereas the optimization problem in Eq. (4) is not jointly convex. So it is commonly solved by computing the coefficients s with fixed D , and computing dictionary D with the coefficients s fixed.

It can be found that none of the SC algorithms has the best sparse reconstruction performance without knowing the prior

information about signal sparsity level, noise strength, etc. Moreover, most recent dictionary learning methods such as K-SVD algorithm learn redundant dictionaries with long computation time, because the whole training dataset should be processed to minimize a constrained objective cost function at a time.^{19,29} Due to the large data samples in vibration signal, it is essential to develop advanced sparse coding and dictionary learning schemes to extract impulse features for machinery fault detection.

3. FSC algorithm and online dictionary learning

3.1. FSC scheme

For sparse coding problem, the performance of solving sparse coefficients depends on appropriate sparse coding algorithm for input signals. In Eq. (3), sparse coding problem with fixed redundant dictionary D is transformed into an l_1 - norm regularized linear least-squares problem, which can be solved using different algorithms such as BP, CD, Lasso, etc. However, it is difficult to determine which SC algorithm can achieve a better sparse reconstruction performance.

Fusion of different sensors data will provide a more robust and accurate estimation.¹⁶ FSC scheme is proposed to fuse different SC algorithms to improve the performance of solving sparse coefficients and it has two main steps. Firstly, several SC algorithms are executed in parallel to estimate sparse coefficients independently. Then, the estimated results are collected and fused to get a new estimate of the sparse coefficients. Any SC algorithm can be used as a participating algorithm in FSC. FSC method has no limit on the number of participating algorithms. In this paper, we assume that $m \geq 2$ denotes the number of different participating SC algorithms, and K is the sparse level which keeps only the most K -dominated sparse coefficients. For the j th participating SC algorithm ($j = 1, 2, \dots, m$), \hat{s}_j denotes the estimated sparse coefficients by current SC algorithm, and \hat{A}_j is defined as a support set that indicates the positions of K -dominate elements in \hat{s}_j . The union of support sets estimated by different SC algorithms is defined as joint support set $\Gamma = \bigcup_{j=1}^m \hat{A}_j$, and Γ^c denotes the complement set of Γ . $|\Gamma|$ denotes the size of joint support set Γ and we assume that $|\Gamma| = q \leq p$. So the problem in Eq. (1) is converted into a low dimensional problem as follows:

$$x = D_\Gamma s_\Gamma + w \quad (5)$$

where $D_\Gamma \in \mathbb{R}^{p \times q}$ and $s_\Gamma \in \mathbb{R}^{q \times 1}$. The pseudo-inverse matrix D_Γ^\dagger can be computed based on the assumption of $q \leq p$. Therefore, we can use least-square approach to solve Eq. (5) to estimate sparse coefficients. The procedure of FSC is shown in Table 1.

In Table 1, \hat{A} denotes the support set estimated by FSC algorithm and $\hat{A} \subset \Gamma$. \hat{A}^c denotes the complement set of \hat{A} which indicates the position of non-dominated elements in \mathbb{V}^K . FSC algorithm solves sparse coefficients using a simple least-squares approach, and the participating algorithm can be any SC algorithm without any modification. In view of the difficulty in selecting appropriate SC algorithm, the union of support set Γ in FSC always contains at least as many dominated coefficients as the support sets estimated from singular best performing SC algorithm. The fusion scheme of SC algorithm can provide improvement on solving sparse

Table 1 Procedure of FSC.

Algorithm 1: FSC

Require: $\mathbf{D} \in \mathbb{R}^{p \times n}$, $\mathbf{x} \in \mathbb{R}^p$, K , and $|\Gamma| \leq p$ **Initialization:** $\mathbf{V} = \mathbf{0} \in \mathbb{R}^p$ **Fusion:****Step 1.** Different participating SC algorithms are executed independently, and the support sets $\{\hat{A}_j\}_{j=1,m}$ are computed**Step 2.** Compute joint support set $\Gamma = \bigcup_{j=1}^m \hat{A}_j$;**Step 3.** $\mathbf{V}_\Gamma = \mathbf{D}_\Gamma^+ \mathbf{x}$, $\mathbf{V}_{\Gamma^c} = \mathbf{0}$;**Step 4.** Let \mathbf{V}^K denote the best K -sparse approximation of \mathbf{V} , and \hat{A} is the position of K -dominated elements in \mathbf{V}^K **Output:** $\hat{\mathbf{s}}_{\hat{A}} = \mathbf{D}_{\hat{A}}^+ \mathbf{x}$, $\hat{\mathbf{s}}_{\hat{A}^c} = \mathbf{0}$, the sparse coefficients estimated by FSC is $\hat{\mathbf{s}} = \hat{\mathbf{s}}_{\hat{A}} \cup \hat{\mathbf{s}}_{\hat{A}^c}$

coefficients and reconstructing sparse signal. However, FSC method has higher computational complexity because different participating algorithms are executed in parallel.

3.2. Improved online dictionary learning method

Vibration signals provide useful information to extract impulse features for machinery fault detection. However, the measured signals usually contain large number of data samples. Classical dictionary learning methods update each column of the current redundant dictionary by using the whole training set and the dictionary obtained in previous iteration, which cause low computational efficiency of training dictionary.

Assuming that the vector $\mathbf{x}_i \in \mathbb{R}^p$ is one sample in training set $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$ and M denotes the number of iterations. The problem of training redundant dictionary \mathbf{D} in Eq. (4) can be solved by using different dictionary learning methods. In classical learning methods, each column of dictionary \mathbf{D}' at current iteration is updated by using the whole training set \mathbf{X} and the dictionary \mathbf{D}^{t-1} obtained at the previous iteration. For example, the iterative procedure of updating dictionary using typical steepest descent learning algorithm can be described as follows:³⁰

$$\mathbf{D}' = \mathbf{D}^{t-1} - \eta \sum_{i=1}^N (\mathbf{D}^{t-1} \mathbf{s}_i - \mathbf{x}_i) \mathbf{x}_i^T \quad (6)$$

where η denotes the parameter of learning rate. To avoid the values of \mathbf{D} to be awfully large, it is common to constrain its columns $\{\mathbf{d}_j\}_{j=1}^n$ to have a l_2 -norm less than or equal to one. The convex set of matrices Ω is subject to the constraint:

$$\Omega = \{\mathbf{D} \in \mathbb{R}^{p \times n} \text{ s.t. } \forall j = 1, 2, \dots, n, \|\mathbf{d}_j\|_2 \leq 1\} \quad (7)$$

Based on stochastic approximations theory, online dictionary learning method is presented to train redundant dictionary with low computational cost and consumption.²⁹ In online method, the samples of training set $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$ are assumed as independently and identically distributed (i.i.d) samples with a probability distribution $P(\mathbf{x})$.¹⁹ Based on stochastic approximation algorithm, the new redundant dictionary \mathbf{D}' at t th iteration can be computed by minimizing the expected cost function $\hat{f}_t(\mathbf{D})$ in Eq. (8)

$$\begin{aligned} \mathbf{D}' &= \arg \min_{\mathbf{D} \in \Omega} \hat{f}_t(\mathbf{D}) \\ &= \arg \min_{\mathbf{D} \in \Omega} \frac{1}{t} \sum_{k=1}^t \frac{1}{2} \|\mathbf{x}^k - \mathbf{D} \mathbf{s}^k\|_2^2 + \mu \|\mathbf{s}^k\|_1 \end{aligned} \quad (8)$$

where \mathbf{x}^k is the sample drawn from probability distribution function $P(\mathbf{x})$ at k th iteration. In practice, the i.i.d samples can be commonly obtained by choosing data sequentially in training set²⁹ or using a Markovian process method.³¹ \mathbf{s}^k the sparse coefficients computed at k th iteration. Instead of using the whole training set \mathbf{X} , online dictionary learning method only uses small training samples drawn from probability distribution function, which can improve the efficiency of training redundant dictionary. However, singular SC algorithm (LARS-Lasso) was used to solve the sparse coefficients in the online learning method,^{19,29} which may not compute the sparse coefficients accurately.

Based on FSC method in Section 3.1, an improved online dictionary learning method is presented to obtain the redundant dictionary with high accuracy. The procedure of the improved online dictionary learning method can be seen in Table 2. The improved method includes sparse coefficients' solving and redundant dictionary updating.

- (1) Sparse coefficients' solving. The sparse coefficient \mathbf{s}^k is used to update the redundant dictionary at each iteration, and it is important to ensure the accuracy of solving sparse coefficients. Thus, FSC scheme is used to solve the sparse coefficients in Step 3 accurately. And the participating SC algorithms can be any SC algorithm such as BP, Lasso, CD, etc.
- (2) Redundant dictionary updating. The redundant dictionary \mathbf{D} can be updated with high training speed based on stochastic approximation algorithm. The key point of Algorithm 2 is that dictionary \mathbf{D}^k can be updated each time by using stochastic approximation algorithm in Step 6. This step only needs small training samples to learn dictionary at each iteration. Basis functions $\{\mathbf{d}_j\}_{j=1}^n$ in redundant dictionary are obtained by using online learning techniques. It has been proved that \mathbf{d}_j^k in Step 6 gives the solution of learning dictionary \mathbf{D}^k in Eq. (8) with updating columns of dictionary sequentially.¹⁹

Generally, our improved online dictionary learning method handles small training samples drawn from distribution function $P(\mathbf{x})$ at each iteration, which can solve the constrained optimization problem effectively. Additionally, the online learning method can be executed without learning rate parameter, which has great effect on the performance of training redundant dictionary. Although the complexity of algorithm increases due to the usage of FSC scheme, the method can still compute the redundant dictionary \mathbf{D} with

Table 2 Procedure of Improved online dictionary learning.**Algorithm 2:** Improved online dictionary learning

Require: $D^0 \in \mathbf{R}^{p \times n}$ (initial dictionary), $X = \{x_i\}_{i=1}^N$ (training set), $x^k \in \mathbf{R}^p \sim P(x)$ (i.i.d sample drawn from distribution P), μ (regularization parameter), M (number of iterations)
Initialization: $A_0 = 0$, $B_0 = 0$ (intermediate variables)
Repeat:
Step 1. for $k = 1$ to M do
Step 2. Draw one sample x^k from probability distribution $P(x)$ using method in Ref.²⁹
Step 3. Sparse coefficients solved using FSC method in Section 3.1
 $s^k = \arg_{s \in \mathbf{R}^n} \min \frac{1}{2} \|x^k - D^{k-1}s\|_2^2 + \mu \|s\|_1$
Step 4. $A_k = A_{k-1} + s^k (s^k)^T$, $B_k = B_{k-1} + x^k (s^k)^T$, $D^{k-1} = [d_1^{k-1}, d_2^{k-1}, \dots, d_n^{k-1}] \in \mathbf{R}^{p \times n}$
Step 5. Update D^k in Eq.(8) based on stochastic approximation algorithm with D^{k-1} as warm restart
Set $A = \sum_{l=1}^k s^l (s^l)^T$, $A = [a_1, a_2, \dots, a_n] \in \mathbf{R}^{n \times n}$, $B = \sum_{l=1}^k x^l (s^l)^T$, $B = [b_1, b_2, \dots, b_n] \in \mathbf{R}^{p \times n}$
Step 6. The columns of D^k at current iteration are updated sequentially
Repeat: for $j = 1$ to n do
Update the j th column of D^k by the following equations
 $d_j^k = \frac{1}{\max(\|u_j\|_2, 1)} u_j$, $u_j = \frac{1}{A_{jj}} (b_j - D^{k-1} a_j) + d_j$
End for
Step 7. **Return:** update dictionary D^k at k th iteration
Step 8. **End for**
Step 9. **Output:** Redundant dictionary D

high accuracy and low consuming time because of processing small training samples. It has superiority to dealing with large dataset of training samples compared with other dictionary learning methods.

4. Impulse feature extraction method using FSC and online dictionary learning

4.1. Impulse components' extracting scheme

Vibration signals often contain different kinds of noises in measurement process, and impulse features are difficult to extract from noisy signals for machinery fault detection. Sparse coding and dictionary learning algorithm can find concise and high-level sparse representations of input signal, which is the key procedure of extracting impulse features.³² Large data samples in measured noisy vibration signal may lead to slow dictionary learning process and low sparse reconstruction accuracy. FSC and online dictionary learning method can overcome the disadvantages of classical methods, which are very suitable to extract impulse features from noisy input signals.

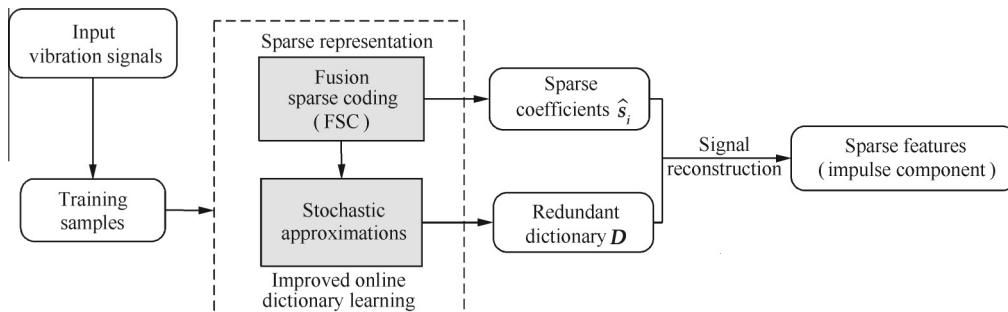
An impulse feature extraction scheme is proposed using FSC and improved online dictionary learning methods (see Fig. 1). The scheme includes two critical procedures: sparse representation and online dictionary learning.

As many dictionary learning problems in processing image or speech signals, the input signals are divided into multi-segments to train dictionary effectively.³³ To reduce the dictionary learning time, vibration signal should also be decomposed into overlapping segments to construct training samples. Assume that $X_0 \in \mathbf{R}^L$ is a one-dimensional vibration signal with large data samples, and $Z \in \mathbf{R}^L$ denotes noiseless signal. The matrix $R_i \in \mathbf{R}^{p \times L}$ ($L \gg p$) is defined as an operator that converts the original signal X_0 into overlapping segments. Samples in training set $X = \{x_i\}_{i=1}^N$ can be described as

$$x_i = R_i X_0 \quad (i = 1, 2, \dots, N) \quad (9)$$

In the step of sparse representation, sparse coefficient \hat{s}_i corresponding to the sample x_i is solved using FSC algorithm. Improved online dictionary learning method can obtain the redundant dictionary adapted to statistical structures of vibration signal. In the step of improved online method, redundant dictionary D is trained using Algorithm 2 with low computation time, and the dictionary D is used to represent vibration signal sparsely. With fixed D and \hat{s}_i , the noiseless signal \hat{Z} can be estimated through solving the optimization problem in Eq. (10).

$$\hat{Z} = \arg \min_Z \lambda \|X_0 - Z\|_2^2 + \sum_{i=1}^N \|D \hat{s}_i - R_i Z\|_2^2 \quad (10)$$

**Fig. 1** Block diagram of the proposed approach for impulse features extraction.

The optimization problem in Eq. (11) is a quadratic term that has a closed-form solution.³⁴

$$\hat{\mathbf{Z}} = \left(\lambda \mathbf{I} + \sum_{i=1}^N \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \left(\lambda \mathbf{X}_0 + \sum_{i=1}^N \mathbf{R}_i^T \mathbf{D} \hat{\mathbf{s}}_i \right) \quad (11)$$

where $\hat{\mathbf{Z}}$ denotes the estimated noiseless signal which is recovered using \mathbf{D} and $\hat{\mathbf{s}}_i$, and the original signal can be represented sparsely using FSC and improved online dictionary learning method. The parameter λ is dependent on the noise level of input signal and small values of λ can achieve better results when the noise level increases.³⁴ Noise standard deviation σ is the indicator of noise level, which is used to compute parameter λ . Based on FSC and improved online dictionary learning method, it is easy to extract impulse features by means of reconstructing sparse signal and eliminating noises from original signal.

4.2. Simulations and comparisons

The impulse components in vibration signals are the most important features of machinery defects, but impulse features are usually contaminated by heavy background noises in measured signals. In order to verify the performance of the proposed method, the simulated vibration signal of rolling element bearing is chosen as the analysis signal. A mathematical simulation model was presented to describe the defects of bearings,¹¹ and the vibration signal \mathbf{X}_0 can be simulated as follows:

$$\begin{cases} \mathbf{X}_0 = \sum_{i=1}^{M_0} A_i \mathbf{S}(t - iT - \tau_i) + w(t) \\ A_i = A_0 \cos(2\pi f_m t + \varphi_A) \\ \mathbf{S}(t) = e^{-Bt} \sin(2\pi f_n t + \varphi_w) \end{cases} \quad (12)$$

where A_i and T are amplitude and period of impulse signal respectively, τ_i the phase of the impulse, and $w(t)$ is the additive zeros-mean white noise with noise level σ . f_m the frequency of amplitude modulator, φ_A the phase of amplitude modulator, f_n the natural frequency related to bearing, φ_w the phase related to bearing, and B the coefficient of resonance damping.

The sample rate is 20 kHz and the length of vibration signal L is 8192. We choose the parameters of impulse amplitude $A_0 = 1$, frequency $f_m = 0.5$ kHz, phase $\tau_i = \varphi_A = \varphi_w = 0$, period $T = 0.02$ s and natural frequency $f_n = 1$ kHz. The coefficient of resonance damping B is set as 100π . Signal-to-noise ratio (SNR) is used to evaluate the standard deviation of noiseless impulse components compared to that of the additive noise. We simulate various vibration signals with different SNRs to further test the performance of the proposed method. Fig. 2(a) is the time-domain waveform of simulated noiseless impulse signal, and Fig. 2(b) shows the time-domain waveform of impulse signal under -10 dB zeros-mean Gaussian white noise.

The scheme of the proposed method in Fig. 1 is used to extract impulse features from simulated noisy vibration signal \mathbf{X}_0 in Fig. 2(b). Due to the large data samples, raw signal \mathbf{X}_0 is locally processed to learn dictionary efficiently. Input signal \mathbf{X}_0 should be firstly divided into segments with data points $p = 64$ to construct the training set \mathbf{X} using operators $\mathbf{R}_i (i = 1, 2, \dots, N)$. Overlapping rate is an important parameter which determines the size of training set.¹² It is commonly used to balance between computational efficiency and training

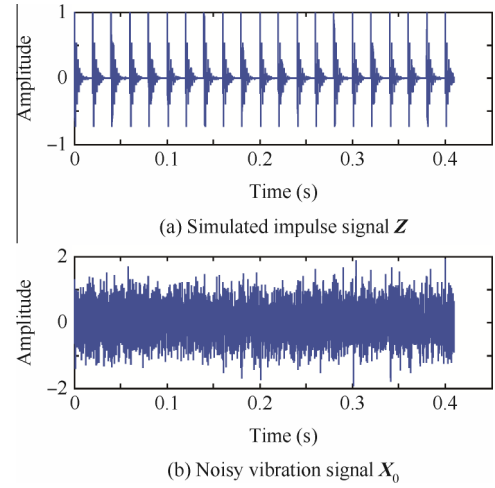


Fig. 2 Time-domain waveform of simulated vibration signal.

accuracy. In this example the parameter is set as overlap ratio of 50%. The redundant dictionary is trained using improved online learning method in Algorithm 2, and the dimension of the initial dictionary \mathbf{D}^0 is set as 64×128 . Classical regularization parameter μ is $1/\sqrt{p}$,³⁵ but the parameter μ is experimentally set as $1.2/\sqrt{p} = 0.15$ in this simulation example which ensures a reasonable sparsity level of data samples (not more than 10 K -dominate coefficients). Different levels of noises are added to original impulse signal, and we find empirically that the parameter $\lambda = 3/\sigma$ can achieve the best accuracy of sparse signal reconstruction with various SNRs. The number of iterations for online dictionary learning algorithm M is 30.

FSC scheme is used to solve sparse coefficients and the participating SC algorithms include BP, Lasso and CD. We also use the same participating SC algorithms in FSC scheme to solve sparse coefficients for improved online dictionary learning in Algorithm 2. To compare the performance of different methods, the impulse components are extracted from raw signal \mathbf{X}^0 using FSC method with the improved online dictionary learning and singular SC method with traditional online method (see Fig. 3). Results of reconstructed impulse components using different methods are assessed by average root mean squared errors (RMSE) as follows:

$$\text{RMSE} = \frac{1}{Q} \sum_{i=1}^Q \sqrt{\frac{1}{L} \sum_{j=1}^L (\mathbf{Z}(j) - \hat{\mathbf{Z}}(j))^2} \quad (13)$$

where L is the length of simulated noisy signal, and $\hat{\mathbf{Z}}$ the impulse components reconstructed by different SC methods. To each level of SNR, ten noisy signals are simulated and the parameter $Q = 10$. The average RMSE between original signal \mathbf{Z} and reconstruction signal $\hat{\mathbf{Z}}$ are listed in Table 3.

It can be seen from Fig. 3(a)–(c) that the additive noise in simulated vibration signal has been reduced greatly. With the values of SNR increasing, FSC scheme has better performances in extracting impulse components than other singular SC methods. The corresponding envelope spectra based on Hilbert transform are shown in Fig. 3(d)–(f). For the simulation model of defective bearing in simulated vibration signal \mathbf{X}_0 , the impulse characteristic frequency $f = 1/$

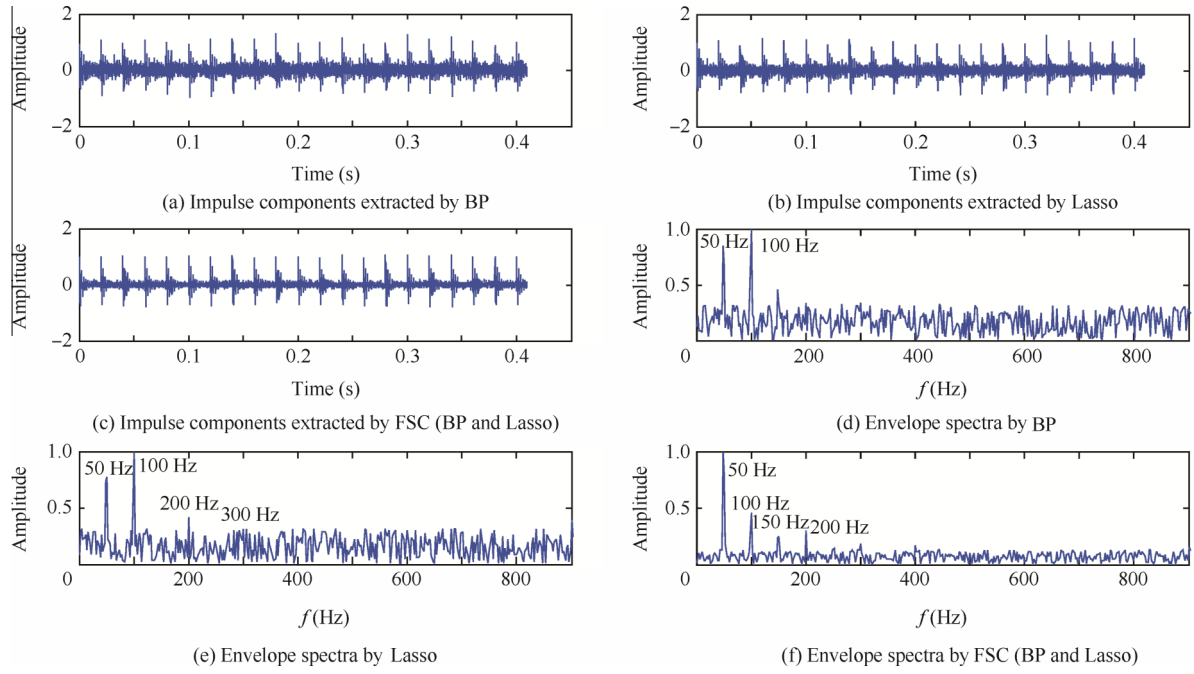


Fig. 3 Time-domain waveforms of impulse components extracted by different SC algorithms.

Table 3 Comparison of average RMSE with various SNRs using different sparse coding methods.

Method	SNR					
	-10 dB ($\sigma \approx 0.47$)	-6 dB ($\sigma \approx 0.29$)	-2 dB ($\sigma \approx 0.19$)	0 dB ($\sigma \approx 0.15$)	2 dB ($\sigma \approx 0.12$)	4 dB ($\sigma \approx 0.09$)
BP	0.122	0.104	0.097	0.080	0.068	0.047
Lasso	0.107	0.093	0.086	0.069	0.055	0.036
CD	0.101	0.091	0.072	0.063	0.050	0.033
FSC(BP, Lasso)	0.089	0.078	0.067	0.059	0.057	0.044
FSC(BP, CD)	0.084	0.071	0.062	0.052	0.045	0.028
FSC(Lasso, CD)	0.079	0.068	0.057	0.042	0.037	0.025
FSC(BP, CD, Lasso)	0.072	0.066	0.057	0.039	0.036	0.022

$T = 50$ Hz and its harmonics (100, 150, 200) Hz can be markedly identified using FSC method in Fig. 3(f). However, due to the heavy background noises, it is difficult to distinguish whether the impulse characteristic frequency is 50 or 100 Hz using BP or Lasso algorithm in Fig. 3(d)–(e). Especially in Fig. 3(e), with the large amplitude of the frequency in 100 and 200 Hz, the impulse characteristic frequency will be wrongly regarded as 100 Hz with its harmonic components (200, 300) Hz. In Fig. 3(f), FSC method can identify 50 Hz impulse characteristic frequency accurately compared with other singular SC methods. From Table 3, FSC(BP, CD, Lasso) method has the least RMSE and it can significantly improve the sparse reconstruction performances. As the additive noise level decreases, impulse components are extracted correctly and RMSE also decreases using different methods. Grey part in Table 3 indicates that FSC method may not be the optimal solutions, because the jointly support sets estimated by FSC(BP, Lasso) are not superior to that of CD algorithm. However, without prior information of original input signal, FSC method still has better performance in extracting impulse features correctly than other singular SC methods.

The training set \mathbf{X} consists of 64×255 samples. We use different dictionary learning methods to train redundant dictionary \mathbf{D} and compute the time consumed. These redundant dictionaries are produced using K-SVD and the improved online dictionary learning method (executed 10–100 iterations with 10 interval, using BP and FSC(BP, Lasso) for sparse coding). All the simulations in this paper are running on a dual-core 2.93 GHz CPU machine with 2 GB RAM using MATLAB2009A implementation. In Fig. 4, we compare the consumed time of K-SVD and the improved online dictionary learning methods using FSC(BP, Lasso) algorithm.

The simulation shows that the improved online method executes significantly faster than K-SVD in training redundant dictionary \mathbf{D} . K-SVD method is essentially a second-order iterative batch procedure, which uses the whole training set to minimize the objective cost function at each iteration; whereas online method only processes one i.i.d sample drawn from the training set \mathbf{X} with distribution $P(\mathbf{x})$ at a time and it can solve the dictionary learning problem more efficiently.

To further evaluate the performance of the proposed method in impulse components extraction, the wavelet

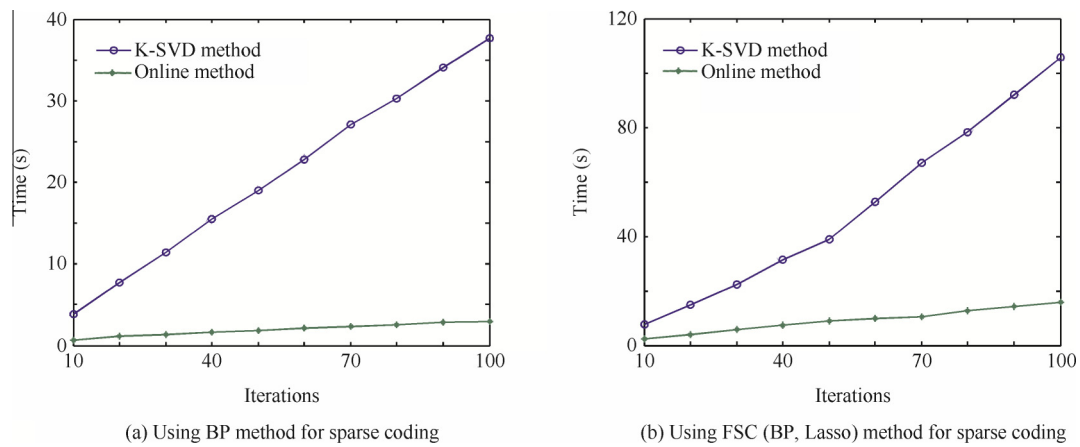


Fig. 4 Comparisons of consumed time by using different dictionary learning methods.

Table 4 Comparison of average RMSE and computation time by different methods.

Method	Average RMSE	Running time(s)
Wavelet shrinkage	0.1497	2.141
BP and K-SVD (30 iterations)	0.1214	13.673
FSC(BP, CD) and improved online learning (30 iterations)	0.0840	7.465

shrinkage method⁴ and basic pursuit de-noising (BPDN) method³⁶ are used to process the simulated heavy noisy vibration signal with SNR = -10 dB for comparison. The noise level σ is important for signal analysis and processing, which can be estimated as $\hat{\sigma} = \text{median}(|w_1 - \text{median}(w_1)|)/0.6745$ based on the wavelet coefficients of input signal at the finest scale resolution level,³⁷ where w_1 denotes the orthogonal wavelet coefficients at the finest scale and in this example noise level $\sigma \approx 0.065$. The threshold of wavelet shrinkage method is set as $\sigma\sqrt{2\ln(L)/L}$,⁴ where L is the length of vibration data. We use wavelet shrinkage method to extract impulse components with the above chosen threshold. BPDN is also used to process the simulated signal for comparison and the redundant dictionary is trained based on K-SVD method. Ten noisy signals are simulated independently to evaluate the performances of different methods and the parameters $Q = 10$. The average RMSE between the extracted impulse components and simulated noiseless signal are computed and the average running time with different methods is also compared in Table 4.

Table 4 shows that FSC and improved online dictionary learning method has the least RMSE and the average running time is reasonable. Compared with other impulse feature extraction methods, the proposed method can reconstruct sparse signal and extract impulse components embedded in heavy noisy vibration signal accurately and effectively.

5. Experimental example

In order to validate the effectiveness and performance of the proposed method, a practical example of extracting fault features from aircraft engine rotor vibration signal with heavy background noises is presented. The proposed method is used to process the vibration signal and extract impulse features compared with wavelet shrinkage method and the BPDN method.

The vibration data is acquired from the aircraft engine rotor experimental platform provided by Nanjing University of Aeronautics and Astronautics. The platform consists of the rotor, a spindle driving motor, rolling bearings, pedestal mount and couplings. The vibration data acquisition system and platform installation sketch can be seen in Fig. 5. Both vertical and horizontal vibration signals are collected using acceleration sensors at 20 kHz sample frequency. The rotational speed of the spindle driving motor is 1500 r/min.

Bearings are the critical parts of the aircraft engine rotor system and it is important to extract the fault features from vibration data for classifying the fault types of bearings. Thus, the bearing run-to-failure test is carried out to obtain the fault data in this experimental platform. The vibration data is acquired from sensors attached to the test bearing in both vertical and horizontal directions. In order to acquire the fault data rapidly, we use the bearing with outer race fault at the beginning of the run-to-failure test. The test bearing used in this experiment is deep groove ball bearing 6309E and the

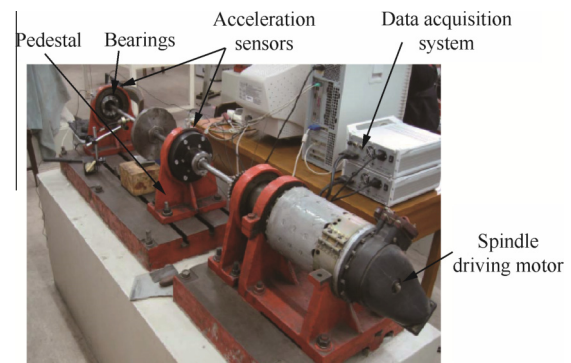


Fig. 5 Experimental platform of aircraft engine rotor.

outer race fault characteristic frequency f_i can be computed as follows:¹

$$f_i = \frac{N_b f_r}{2} \left(1 - \frac{d_b \cos(\theta)}{d_p} \right) \quad (14)$$

where N_b denotes the number of balls, f_r the rotational frequency of the rotor, d_b the ball diameter, d_p the ball pitch diameter and θ the ball contact angle. The test bearing's parameters and the outer race fault characteristic frequency f_i are shown in Table 5.

A group of measured vibration data collected from one vertical acceleration sensor is shown in Fig. 6(a). The measured raw signal has 10,000 data points, and impulse features are mainly buried with heavy background noises. As can be seen from the envelope spectrum in Fig. 6(b), the rotational frequency $f_r = 25$ Hz is presented. But the fault characteristic frequency $f_i = 60$ Hz cannot be identified clearly due to the heavy noises. Therefore, the proposed method is used to extract impulse features from original vibration signal for weak fault detection.

For the purpose of evaluating the performances of impulse features extraction, we use the proposed method to process the measured vibration signal compared with wavelet shrinkage and BPDN methods. The noise level σ can be estimated by using the method in Section 4.2, and in this example noise level $\sigma \approx 0.05$. The threshold of wavelet shrinkage method is set as $\sigma \sqrt{2 \ln(L)/L}$. The above chosen threshold is used in wavelet shrinkage method to extract impulse components. BPDN is also used to process the vibration signal for comparison and the redundant dictionary is trained based on K-SVD method. The impulse features extracted by using wavelet shrinkage and BPDN method can be seen in Fig. 7(a) and (b). Fig. 7(d)–(e) are the corresponding envelope spectra of signal waveforms.

In our method, the raw vibration signal is divided into small segments to construct training samples. To obtain the redundant dictionary which can represent the original

vibration signal sparsely with low computational time, the segment size and overlap size are set as 50 and 25 respectively. For improved online dictionary learning method, the dimension of the initial dictionary D_0 is 50×100 , regularization parameter u is experimentally set as 0.14 (not more than 8 K -dominated coefficients) with a reasonable sparsity level. We test several values for parameter λ and the best result is obtained with $\lambda = 0.1/\sigma = 2$. The number of iterations M is 30 in this algorithm. BP and Lasso algorithm are the participating sparse coding algorithms in FSC method. The reconstructed impulse signal and its corresponding envelope spectrum using the proposed method are shown in Fig. 7(c) and (f) respectively.

Fig. 7 displays the waveforms of reconstructed impulse components and the corresponding envelope spectra using different de-noising methods. Wavelet shrinkage method removes much energy in the process of reducing noise. As can be seen from the envelop spectrum in Fig. 7(d), the amplitude of each characteristic frequency is small due to the large energy loss of raw signal, and the triple harmonics ($3f_i$) cannot be distinguished because of the low amplitude. In some cases, wavelet shrinkage method may discard some important features from the measured vibration signal. BPDN method uses BP algorithm as SC algorithm with K-SVD dictionary learning. BP algorithm does not have a good sparse reconstruction performance because of the heavy background noises, and the impulse characteristic frequency is not identified correctly in Fig. 7(e). However, the noises in measured vibration signal have been removed evidently in Fig. 7(c). Due to the influence of bearing outer race fault, the impulse characteristic frequency $f_i = 60$ Hz and its harmonics ($2f_i$, $3f_i$) are remarkable features in envelop spectrum, which can be detected correctly in Fig. 7(f). The impulse features can be extracted from heavy noisy vibration signal using FSC and online dictionary learning method. Table 6 compares the computation time for training redundant dictionary with

Table 5 Test bearing's parameters and fault characteristic frequency.

Test bearing's parameter					Characteristic frequency
N_b	f_r (Hz)	d_b (mm)	d_p (mm)	θ (°)	f_i (Hz)
8	25	40	100	0	60

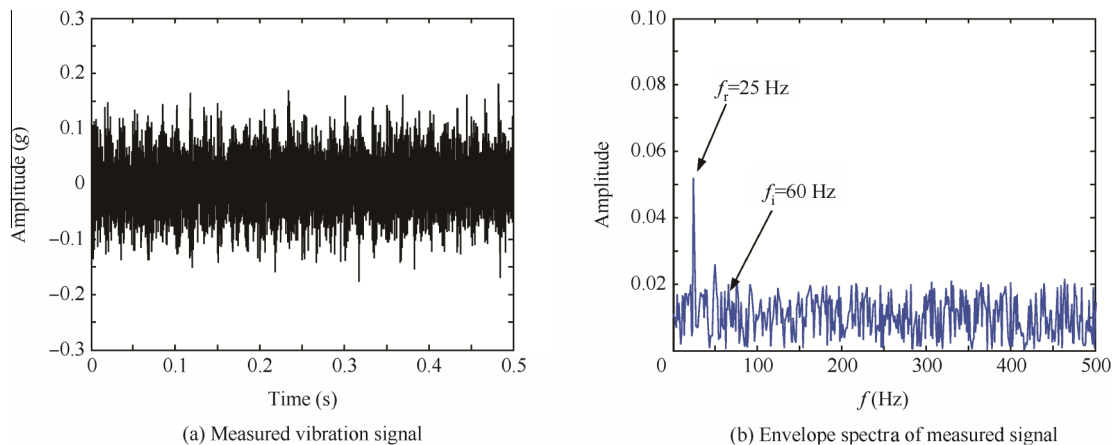


Fig. 6 Measured vibration signal from acceleration sensor.

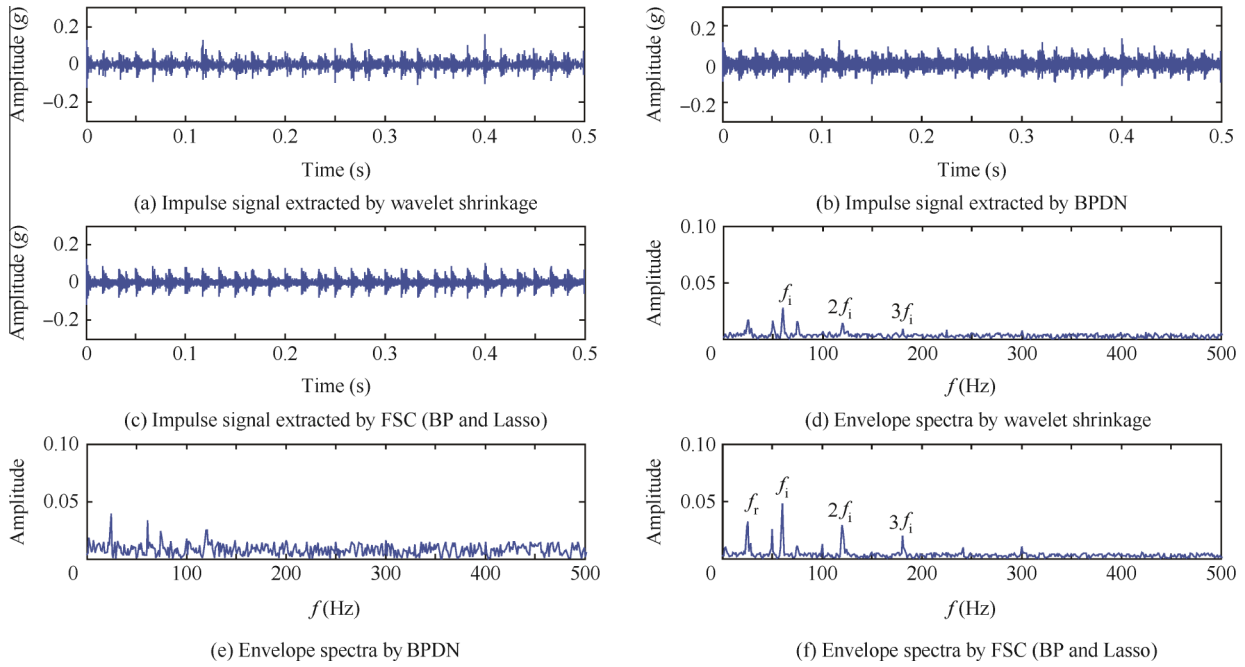


Fig. 7 Time-domain waveforms of impulse features extracted using different de-noising algorithms.

Table 6 Comparison of computation time with various training set sizes using different de-noised methods.

Method	Training set size (30 iterations)		
	50 × 100	80 × 100	100 × 100
BP and K-SVD(s)	3.122	8.451	11.231
FSC(BP, Lasso) and improved online learning(s)	3.874	6.966	10.525
FSC(Lasso, CD) and improved online learning(s)	4.245	8.172	11.651

various training set sizes using different methods. All the experiments are carried out on a dual-core 2.93 GHz CPU machine with 2 GB RAM using Matlab2009A implementation.

It can be seen from Table 6 that the running time of the proposed method is reasonable, especially in dealing with large training datasets. Although several participating sparse coding algorithms in FSC method are executed in parallel, the computation time of our method is still lower than K-SVD method because the improved online learning method trains redundant dictionary rapidly using small training samples.

The experiment demonstrates that our method can keep most of energy in raw signal and identify impulse feature frequency correctly. Compared with other de-noising algorithms, the proposed method has good performances on extracting impulse features from heavy noisy signal with reasonable computation time, which can be used for machinery fault feature extraction and incipient fault detection.

6. Conclusions

- (1) In this paper, we propose an impulse feature extraction method based on fusion sparse coding and improved online dictionary learning. Fusion sparse coding

algorithm achieves better sparse reconstruction accuracy than any singular sparse coding algorithm, which is also used to improve the performance of traditional online dictionary learning method. The improved online dictionary learning method can obtain redundant dictionary by using small training samples with high accuracy and reasonable computation time.

- (2) The vibration signal can be reconstructed in sparse domain using FSC and the improved online dictionary learning method, which is the key issue of extracting impulse features for fault detection. The simulation validates that the proposed method can achieve good performance in reducing noises and extracting impulse components from raw signal, especially for the vibration signal with heavy background noises.
- (3) The application to processing aircraft engine rotor vibration signal indicates that this method can be generalized into impulse features extraction and weak fault detection. Compared with other de-noising methods, this method can extract impulse features and detect the fault feature frequency correctly and efficiently.
- (4) Deeper research about selecting the parameters of segment size, overlapping rate and training samples length needs to be further studied.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 51201182). We would like to thank the authors of Refs.^{18,19} for sharing the codes and Mr. Qiang of Nanjing University of Aeronautics and Astronautics for providing experimental vibration data of aircraft engine rotor.

References

- Li RY, Sopon P, He D. Fault features extraction for bearing prognostics. *J Intell Manuf* 2012;**23**(2):313–21.
- Yan RQ, Gao RX. Energy-based feature extraction for defect diagnosis in rotary machines. *IEEE Trans Instrum Meas* 2009;**58**(9):3130–9.
- Gary YG, Lin KC. Wavelet packet feature extraction for vibration monitoring. *IEEE Trans Industr Electron* 2000;**47**(3):650–67.
- Donoho DL. De-noising by soft-thresholding. *IEEE Trans Inf Theory* 1995;**41**(3):613–27.
- Qiu H, Lee J, Lin J, Yu G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J Sound Vib* 2006;**289**(4):1066–90.
- He W, Jiang ZN, Qin Q. A joint adaptive wavelet filter and morphological signal processing method for weak mechanical impulse extraction. *J Mech Sci Technol* 2010;**24**(8):1709–16.
- Yu NN, Qiu TS, Bi F, Wang AQ. Image features extraction and fusion based on joint sparse representation. *IEEE J Sel Top Signal Process* 2011;**5**(5):1074–82.
- Jafari MG, Plumbley MD. Fast dictionary learning for sparse representations of speech signal. *IEEE J Sel Top Signal Process* 2011;**5**(5):1025–31.
- Donoho DL. Compressed sensing. *IEEE Trans Inf Theory* 2006;**52**(4):1289–306.
- Liu HN, Liu CL, Huang YX. Adaptive feature extraction using sparse coding form machinery fault diagnosis. *Mech Syst Signal Process* 2011;**25**(2):558–74.
- Tang HF, Chen J, Dong GM. Sparse representation based on latent components analysis for machinery weak fault detection. *Mech Syst Signal Process* 2014;**28**(1):158–74.
- Chen XF, Du ZH, Li JM, Li X, Zhang H. Compressed sensing based on dictionary learning for extracting impulse components. *Signal Process* 2014;**96**(1):94–109.
- Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process* 2006;**54**(11):4311–22.
- Rubinstein R, Zibulevsky M, Elad M. Double sparsity: learning sparse dictionaries for sparse signal approximation. *IEEE Trans Signal Process* 2010;**58**(3):1553–64.
- Engan K, Aase SO, Husoy JH. Method of optimal directions for frame design. *IEEE international conference on acoustics, speech, and signal processing*, 1999 Mar 15–19 Phoenix, AZ, UK. 1999. p. 2443–6.
- Xu LJ, Zhang JQ, Yan Y. A Wavelet-based multi-sensor data fusion algorithm. *IEEE Trans Instrum Meas* 2004;**53**(6):1539–45.
- Elad M, Yavneh I. A plurality of sparse representations is better than the sparsest one alone. *IEEE Trans Inf Theory* 2009;**55**(10):4701–14.
- Ambat SK, Chatterjee S, Hari KV. Fusion of algorithms for compressed sensing. *IEEE Trans Signal Process* 2013;**61**(14):3699–704.
- Mairal J, Ponce J, Sapiro G. Online learning for matrix factorization and sparse coding. *J Mach Learn Res* 2010;**11**(2):19–60.
- Lewicki MS, Sejnowski TJ. Learning overcomplete representations. *Neural Comput* 2000;**12**(2):337–65.
- Liu JY, Zhu JB. High range resolution profile automatic target recognition using sparse representation. *Chin J Aeronaut* 2010;**23**(5):556–62.
- Olshausen BA, Field DJ. Natural image statistics and efficient coding. *Netw Comput Neural Syst* 1996;**7**(2):333–9.
- Chen SB, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit. *SIAM J Sci Comput* 1998;**20**(1):33–61.
- Tibshirani R. Regression shrinkage and selection via the Lasso. *J Roy Stat Soc B* 1996;**67**(1):267–88.
- Friedman J, Hastie T, Hoefling H, Tibshirani R. Pathwise coordinate optimization. *Ann Appl Stat* 2007;**11**(9):302–32.
- Gorodnitsky IF, Rao BD. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted norm minimization algorithm. *IEEE Trans Signal Process* 1997;**45**(3):600–16.
- Cai T, Wang L, Xu GW. Shifting inequality and recovery of sparse signals. *IEEE Trans Inf Theory* 2010;**56**(9):4388–94.
- Varadarajan B, Khudanpur S, Tran TD. Stepwise optimal subspace pursuit for improving sparse recovery. *IEEE Signal Process Lett* 2011;**18**(1):27–30.
- Bottou L, Bousquet O. The trade-offs of large scale learning. *Adv Neural Inf Process Syst* 2008;**20**(2):161–8.
- Olshausen BA, Field BJ. Sparse coding with an overcomplete basis set: a strategy employed by v1? *Vision Res* 1997;**37**(12):3311–25.
- Benveniste A, Metivier M, Priouret P. *Adaptive algorithms and stochastic approximations*. Berlin: Springer Publishing Company, Incorporated; 1990. p. 141–3.
- Donoho DL, Elad M, Temlyakov VN. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans Inf Theory* 2006;**52**(2):6–18.
- Protter M, Elad M. Image sequence denoising via sparse and redundant representations. *IEEE Trans Image Process* 2009;**18**(1):27–35.
- Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionary. *IEEE Trans Image Process* 2006;**15**(12):3736–45.
- Bickel PJ, Ritov Y, Tsybakov AB. Simultaneous analysis of Lasso and Dantzig selector. *Ann Stat* 2009;**37**(4):1705–32.
- Yang H, Mathew J, Ma L. Fault diagnosis of rolling element bearings using basis pursuit. *Mech Syst Signal Process* 2005;**19**(2):341–56.
- Donoho DL, Johnstone IM. Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 1994;**81**(3):425–55.

Deng Sen received B.S. and M.S. degrees from Air Force Engineering University in 2008 and 2011, respectively, and then became a Ph.D. candidate in the same university. His main research interests are signal processing and fault diagnosis.

Jing Bo received M.S. degree from Air Force Engineering University in 1996, Ph.D. from Northwestern Polytechnical University in 2002, and she is now a professor in Air Force Engineering University. Her current research focuses on prognostics and health management, design for testability, sensor network and information fusion.

Sheng Sheng received B.S. and M.S. degrees from Air Force Engineering University in 2008 and 2011, respectively, and then became a Ph.D. candidate in the same university. His main research interests are prognostics and health management (PHM) and fault diagnosis.